DOCUMENT RESUME

ED 208 054 TM 810 810

AUTHOR Caro, Francis G.
TITLE Leverage and Evaluation Effectiveness.
INSTITUTION Community Service Society of New York, N.Y.
SPONS AGENCY Robert Sterling Clark Foundation, Inc., New York, N.Y.
PUB DATE May 80
NOTE 23p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Evaluators; Financial Support; *Models; *Program Evaluation; *Social Services
IDENTIFIERS *Evaluation Research; Regulatory Agencies; *Regulatory Evaluation

ABSTRACT

Weakness in evaluations often can be traced to structural limitations in the positions of evaluation researchers. Conventional human relations techniques often are an insufficient basis for securing strong support for evaluation research. Strategies for increasing evaluation research leverage are reviewed. Alignment of evaluation research with regulatory bodies with authority to suspend public program expenditures is advocated. Several likely obstacles in the development of the regulatory evaluation model are anticipated and addressed. (Author)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED208054

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

# LEVERAGE AND EVALUATION EFFECTIVENESS*

by

Francis G. Caro **

INSTITUTE FOR SOCIAL WELFARE RESEARCH
COMMUNITY SERVICE SOCIETY
105 East 22nd Street
New York, N.Y. 10010

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

F. G. Caro

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 810

May 1980

* The paper is based in part on research supported by the Robert Sterling Clark Foundation.

** Dwight Frankfather and Berkeley Rice contributed useful suggestions.

HERBERT P. PATTERSON, *Chairman of the Board* JACK JOHN OLIVERO, *President* MRS. WILLIAM A. M. BURDEN, DOUGLAS WILLIAMS, *Vice Chairmen of the Board*
MRS. DONALD GLICKMAN, *Treasurer* SAMUEL J. DUBOFF, *Secretary* BERTRAM M. BECK, *General Director*

ABSTRACT

Weaknesses in evaluations often can be traced to structural limitations in the positions of evaluation researchers. Conventional human relations techniques often are an insufficient basis for securing strong support for evaluation research. Strategies for increasing evaluation research leverage are reviewed. Alignment of evaluation research with regulatory bodies with authority to suspend public program expenditures is advocated. Several likely obstacles in the development of the regulatory evaluation model are anticipated and addressed.

## LEVERAGE AND EVALUATION EFFECTIVENESS

Mere presence in a social program domain does not assure evaluation researchers of the support they require to contribute effectively. Reluctance of program operators to specify objectives, to agree to random assignment of potential clients to control groups, to permit systematic observation of service transactions, to allow thorough, independent pre- and postmeasurement of service recipients on dependent variables, to attend seriously to the implications of evaluation findings, and to assent to publication of disappointing findings are among the obstacles which evaluation researchers frequently encounter. In some instances, of course, evaluators enjoy very strong support and face few if any of the problems listed above. In other instances evaluators experience all of these obstacles and more.

Experienced evaluation researchers are accustomed to dealing with adversity. They make judicious guesses about program intentions and learn to live with criticism for addressing the wrong questions. They regularly use quasi-experimental and even pre-experimental designs, address process variables which are poor substitutes for measures of outcome variables, and make use of flawed data available through service records. In making these methodological compromises, evaluators ultimately invite critical comments from their colleagues (see, for example, Bernstein & Freeman, 1975 and Cook & Gruder, 1978).

At issue, of course, is not simply the evaluator's interest in carrying out sound research but the public interest in effective social

programs. Assuming at worst that social programs are harmless, the public has reason to be concerned about the tax implications of publicly supported programs which continue to be justified on faith rather than evidence. Of potential concern are not only programs receiving direct public funds but privately financed endeavors dependent on their tax exempt status. While federally funded demonstration programs now regularly receive evaluation attention, many ongoing programs continue to receive no evaluation attention at all. The public interested in efficient use of resources, therefore, has reason to be concerned not only about evaluations which are equivocal because of their soft methodology but about the programs entirely untouched by evaluation research.

What can be done to extend the evaluation research domain and increase its leverage? Evaluators are accustomed to finding themselves the advocates not only for powerful methodology but also for evaluation itself. As advocates they typically try to educate and persuade. They rely extensively on human relations techniques to establish their importance, to sustain the interest and cooperation of program personnel, and to persuade clients to attend to evaluation results (see, for example, Caro, 1977).

Clearly some evaluators are highly successful in making a case for evaluation and in persuading funding agencies and program operators to provide the support necessary for powerful evaluations. In estimating what evaluators can expect to accomplish through education, advocacy, and human relations, it is useful, however, to examine the perspectives of other key actors in the social program domain. Because of the

importance of finances and control over operations, funding agencies and administrators play particularly important roles in determining the fate of evaluation concerns. In principle, both those who allocate resources and those who administer programs have reason to support program evaluation. They should be committed to efficient and effective use of scarce resources. They should appreciate the contribution which evaluation can make to program development. The organizational literature and the experience of evaluators, however, suggest that other forces may dampen the enthusiasm of funding agencies and administrators for evaluation.

Because administrators have received more attention in the literature, it is convenient to consider their perspective on evaluation first. Some years ago Etzioni (1960) made the useful observation that organizations are not simply concerned with realization of program objectives but such other matters as organizational survival. Public objectives are sometimes less important than unpublicized organizational goals. Administrators do make commitments to programs for reasons which are not made public, e.g., their own career ambitions, cultivation of important external support, and loyalty to staff. Evalutions addressed to official objectives in some instances might not only embarrass the organization by showing modest results but invite unwelcome questions about unofficial reasons for commitment to the program. As Schulberg & Baker (1968) point out, administrators for these reasons often are careful in identifying programs for which they invite evaluation attention.

Although contraction in public funding for social programs is cited commonly as an argument for expanded emphasis on program evaluation,

scarce resources can contribute to the forces which undermine evaluation. Publicly funded social programs typically are addressed to tough residual problems. To secure resources in a highly competitive market, administrators have learned that it can be useful to project a "miracle worker" image. They learn to convey the conviction that with modest resources they will achieve dramatic results. As Campbell (1969) puts it, "Specific reforms are advocated as though they were certain to be successful." Skillful operators have learned how to use evidence of early, apparently promising results to proclaim success and to command additional resources. Sometimes these administrators welcome the presence of evaluation researchers as a means of enhancing their credibility and prestige. The evaluators, of course, are welcome only to conduct studies which do not challenge fundamental program premises. In Campbell's terms, these administrators are "trapped" by their exaggerated commitments and cannot afford an honest evaluation.

Similarly, funding agencies whether public or private have reasons for ambivalence about evaluation. In issuing grants and contracts, funding bodies are well advised to scrutinize applicants carefully. In principle, funding agencies should seek evidence of effectiveness in program performance as a guide to continuing funding decisions. Funding agencies, however, may be trapped in much the same way as administrators. Seeking to maximize what they can accomplish with scarce resources, funding bodies are attracted to those who promise to accomplish a great deal at a modest cost. Program sponsors, therefore, are highly vulnerable to being victimized by over-advocacy. Sponsors may be able to afford honest evaluations exposing serious limitations

in a few of their programs, but evidence of pervasive weaknesses in supported programs would erode their own credibility. To protect the public confidence they enjoy, funding bodies have reason to be careful about the evaluations they encourage. To the extent to which funding bodies attempt to accomplish a great deal with meager resources, they are increasingly vulnerable to being embarrassed by thorough evaluations.

An additional force which inhibits efforts to establish a constituency for evaluation research might be described as a longing for faith. Not only among program sponsors and program personnel but among clients, legislative bodies, and the general public, there is a desire to believe that problems can be solved, that certain interventions work. On some matters various publics are willing and even eager to be skeptical. Challenges to certain fundamental assumptions, however, are not fully welcome because they are unsettling. The public wants to believe, for example, that education is beneficial and that physicians can effectively treat illness. Evaluation research is a part of a large set of cultural forces which seek increasing rationalization of society. A great deal has been accomplished over a period of several centuries in "advanced societies" in gaining acceptance for challenges to traditional practices. Yet it is important to recognize that the quest for faith remains alive. In some sectors evaluation efforts will continue to attract an indifferent or even somewhat hostile response because some portion of the concerned public is not prepared to have its faith in a program intervention challenged.

## Increased Leverage for Evaluation

In light of the strength of the forces constraining significant evaluation contributions, more than education and advocacy are needed if evaluation is to become a strong presence throughout the social programming domain. A number of models for providing increased leverage for evaluation deserve attention.

Watchdogs and Gadflies. Organizations with a mandate to protect the the public interest increasingly show signs of interest in evaluation research. Outside of formal lines of authority, organizations like the League of Women Voters, Common Cause, and "Nader's Raiders" sponsor inquiries into various public programs. These "watchdogs" use their prestige, the content of their message, and persuasion to influence policy. Independent investigations of the operations of public programs are a well established tradition in American social reform efforts but their explicit link to evaluation research is relatively new. The watchdog who operates out of the private sector is typically constrained greatly by modest financial resources. In light of the more adequate funding of public watchdog agencies, their recent interest in evaluation research is particularly encouraging. Traditionally such units limited themselves to financial accounting. On a federal level, the General Accounting Office which was created to serve Congress increasingly conducts inquiries concerned with program effectiveness. The New York City Comptroller's occasional studies of program performance indicates that this broader conception of public accounting responsibilities is not entirely limited to the federal level. (See, for example, N.Y.C. Office of the Comptroller, 1978.)

For the most part the watchdog must be satisfied to use available data or conduct surveys. Outside of and antagonistic to program authority structures, the watchdog usually cannot conduct true experiments. (Cleverly designed experimental evaluations, however, have been conducted by watchdog groups to test the effectiveness of programs concerned with discrimination in such areas as housing and public accommodations. (See, for example, Wienk et al., 1979.)

A variation on the watchdog approach is the "gadfly" approach. Working alone and without official sanction, the gadfly is opportunistic in gatering data. Sometimes needed information is in the public domain. Frequently the gadfly gains entry to an organization by concealing his full agenda. Disgruntled lower echelon staff members are often key sources of data. Because he has to make use of opportunities as they present themselves, the gadfly may have to be content with qualitative data.

Both the watchdog and gadfly models are to be encouraged as means of calling attention to matters otherwise inaccessible to evaluators. The weakness of both models, however, are conspicuous. Limited opportunities for true experimentation, uncertain access to data, and uncertain influence over decision-making mean that additional models are needed to extend the evaluation domain.

Evaluation Imperialism. Another possibility is for evaluators to seek control over program operations. If the evaluation researcher becomes the program administrator, he may be able to use his authority to decide that evaluations addressed to central issues and employing powerful methodologies are to be conducted. The evaluator-program

administrator also may be able to arrange a budget which assures sufficient funding for strong evaluation. This approach might be termed "evaluation imperialism." As a strong advocate of this approach, Tornatzky (1979) cites the achievements of George Fairweather (1964, 1969). Working in the Veterans Administration, Fairweather achieved administrative control over both a ward for mental patients and a complementary community facility. Not only did Fairweather use his authority to introduce innovative programming but he conducted randomized experiments. Tornatzky points out that Fairweather even used his leverage as service administrator to shift assignments of ward staff to eliminate personnel as a plausible explanation of differences between experimental and control groups. Tornatzky argues that evaluators are often too quick to accept a subordinate role. If they are enterprising and resourceful, evaluators can acquire the authority over programming which may be needed if they are to be able to conduct powerful experimental evaluations.

The imperialist model makes an important contribution in calling attention to the importance of authority. The evaluator who controls decisions about what is to be evaluated and what methods may be employed is in a much better position to conduct evaluations which are substantively significant and methodologically strong than the evaluator who must rely on education and persuasion. Further, the evaluator-program administrator presumably is guided by an interest in using evaluation results to improve program operations.

Evaluation imperialism, however, is not without its limits as a model for increasing evaluation research leverage. Competition for

control over social program operations often is substantial. In spite of their vigorous efforts, evaluation researchers frequently will be unable to obtain authority over program operations. Further, those who do manage to achieve control over programs will find themselves under pressures detracting from their intentions to conduct experimental evaluations. In their classic article on researcher-practitioner relations, Rodman and Kolodny (1964) effectively argued that an inherent strain divides the two roles. Those who attempt to combine the roles find it difficult to reconcile the two sets of responsibilities. Faced with a board, practitioners, and perhaps client advocates who are opposed to random assignment of clients to control groups, the program administrator-evaluator may find maintenance of good relations with key groups in his operational domain more important than true experimentation. Confronted by financial constraints, he may have to compromise with his intention to invest significantly in evaluation research. Perhaps most importantly, to gain and maintain administrative authority and to be able to generate external financial support, the evaluation imperialist is likely to be trapped by over-advocacy in much the same fashion as other administrators. While evaluation imperialism can be useful in extending the evaluation domain, it is not sufficient as a general means of assuring evaluation the leverage it deserves.

Regulatory Evaluation. A third approach which aligns evaluation with regulation is likely to prove of greatest significance in extending and strengthening the evaluation research domain. Evidence of effectiveness would be required as a condition for continued

public funding of social programs. Some evidence of effectiveness through program evaluation even might be required of privately funded programs which want to enjoy tax exempt status. The evaluator-regulator would have authority to suspend public expenditures for programs and tax exempt status for organizations on the basis of lack of evidence of effectiveness.

The proposed regulatory model would extend approaches already initiated. In 1965 the Elementary and Secondary Education Act included a requirement that projects funded through the program be evaluated. Federal agencies financing demonstration programs now routinely require the inclusion of an evaluation component. Some funding agencies go a step further in mandating that evaluations be conducted by an agent independent of the program operation. Presumably the independent evaluator enjoys greater leverage in establishing an evaluation agenda, carrying out the evaluation, and publicizing results than evaluators subordinate to a program operator.

The mandatory evaluation approach falls short of the regulatory model proposed here in that it simply requires that evaluations be conducted. Agencies sponsoring social programs vary widely in their methodological expectations for the evaluations they require. Further, typically there is no requirement that anyone attend seriously to the evaluation results.

Some precedent for the regulatory model also can be found in accrediting and licensing strategies (Glass, 1971). In some sectors accreditation is a condition for licensing or receipt of public funds. The accreditation approach emphasizes qualifications and facilities.

Accreditation emphasizes potential rather than actual performance. It is the shift to an emphasis on methodologically sound evidence of achievement which would differentiate the proposed regulatory evaluation model from traditional accreditation practice.

Authority to impose significant sanctions on programs for which evidence of effectiveness is lacking is central to the regulatory evaluation model proposed here. Through legislation, public regulatory bodies or publicly sanctioned accreditation agencies would be authorized to suspend or terminate public funding on the basis of negative evaluation findings. If established on a federal level, for example, a regulatory evaluator would have authority to suspend the payment of federal funds to states found to be administering ineffective programs. Further, regulatory evaluation bodies would be provided with substantial structural insulation from routine political pressures. Directors of public regulatory evaluation agencies, for example, might be appointed on a long term basis and be subject to removal from office only for gross misconduct. Similarly the legislative authorization might include a provision calling for funding based on a fixed formula tied to program appropriations in the domain to be evaluated.

Regulatory agencies would set and enforce evaluation research standards in their domain. They would articulate minimum sets of outcome variables, establish performance standards, define the ground rules for acceptable evaluation methodologies, and supervise the execution of evaluation studies. Their authority would go beyond the right to observe and examine records to include conducting experimental studies. In conducting randomized experiments, regulatory evaluators would not have authority to deny entitlements. In the case of capped programs,

however, they would have authority to employ randomization to withhold services from certain marginal service applicants for experimental evaluation purposes.

Some variation in institutional arrangements for conducting regulatory evaluation studies is possible. Regulatory bodies might authorize evaluations conducted entirely by independent evaluation groups. Alternately they might sanction a mixed model in which service organizations would maintain internal evaluation units subject to external audit. The required audits might be conducted directly by accreditation agencies or by authorized independent evaluation groups. As Campbell (1977) suggests, the external evaluator would review the methodological adequacy of measurement procedures, the quality of the data, and the bases upon which inferences can be made regarding program effectiveness. In most instances the outside evaluator would depend entirely on data collected by internal evaluators. The outsider might collect some additional data as a check on or extension of internal data collection procedures. Anticipation of a review by external evaluators would create pressure on service organizations to permit sound internal evaluation work. The threat of loss of public funds for failure to produce sound evaluation data would serve as a strong incentive to service agencies to authorize strong internal evaluation units.

Ultimately regulatory evaluators would be expected to terminate programs found to be ineffective. Failure to find evidence of positive results on key outcome variables would be a sufficient basis for termination. Regulatory evaluators would not act precipitously in closing programs. As a first step, the absence of positive outcome

evidence would be reviewed with program personnel. Opportunities would be provided for refining objectives and program strategies. If important new outcome variables or promising program strategies were introduced, new evaluation studies would be conducted. The continued absence of positive outcomes, then, would lead to a termination action.

## Making Regulatory Evaluation Work

No institutional arrangement automatically provides effective solutions to the problems it was designed to address. Difficulties with the regulatory evaluation model can be anticipated. The extent to which they can be overcome will depend on the skill and industry of those who work with development of the regulatory evaluation model. It is useful to consider how some of the inevitable problems might be addressed.

Weaknesses in evaluation methodology will be a source of objections to regulatory evaluation. Some will argue that while current evaluation methodologies provide a basis for raising critical questions about social program effectiveness, they often do not provide the conclusive evidence desired as a basis for decisions regarding the fate of programs. Some programs, for example, are justified on the basis of long term effects which cannot be tested quickly enough to meet the requirements of short term decision cycles. Other programs are justified on the basis of highly abstract objectives which do not readily lend themselves to measurement. Regulatory agencies would be expected to address these problems on a case by case basis. In some instances

evaluation research requirements might be tempered. In other instances reformulation of goal frameworks might be required. Program proponents would be required to find a sufficient basis for justifying social programs in objectives which are relatively immediate and sufficiently concrete so that they lend themselves to straightforward measurement. While some might be disappointed in the evaluation of programs on the basis of immediate outcomes, regulatory evaluators will be able to argue that immediate outcomes are preferable to structure and process variables as bases for evaluation judgments.

The eagerness of evaluation researchers to conduct true experimen based on randomization surely will create concerns among those with direct service responsibility. Regulatory bodies will be expected to abide by ethical and legal principles protecting the rights of subjects of human experimentation. Because of ethical and legal concerns, regulatory evaluators in some instances will have to be satisfied with less than optimal research designs. As indicated above, regulatory evaluators in the interest of conducting true experiments could not deny individuals their entitlements. They will have authority to insist on an experimental design in some form when service resources are insufficient to address aggregate needs. Regulatory evaluators will have to work with providers to find reasonable ways of reconciling service priorities and experimental evaluation interests.

Corruption of outcome measures will be a serious continuing problem for regulatory evaluators (Campbell, 1977). When they know evaluation criteria, program operators can be expected to redirect their activities to obtain artificially high ratings. To the extent

that corruption takes the form of distorted record keeping, evaluators can try to overcome the problem by monitoring record keeping activities or by developing independent data systems. When evaluation measures cover only a portion of the goal domain, program operators can be expected to concentrate their efforts on the domains which are measured at the expense of those which are not. Regulatory evaluators can contend with this problem by seeking comprehensive outcome measurements or by using sampling strategies which are not announced in advance to measure portions of a broad set of outcome variables.

In the case of new program strategies, regulatory evaluators will have to be judicious about their timing in introducing outcome evaluations. Only after program operators have had sufficient opportunity to solve inevitable start-up problems, will it be desirable to introduce the outcome evaluations which will decide the fate of the innovation. In the start-up period, regulatory activities will be limited to financial audits and analyses of structural arrangements, staffing patterns, and service exchanges. In some instances these reviews will reveal needs for corrective action. In other cases they will provide a basis for early termination decisions (Caro, 1977).

Universal, comprehensive regulatory evaluation will be expensive. In part the high cost of regulatory evaluation will be justified by savings realized through elimination of ineffective programs. In order to justify their budgets regulatory agencies will have to demonstrate their utility. In part their ability to attract sufficient resources will depend on the energy and political skills of proponents. At the same time the agendas of regulatory evaluators inevitably will exceed available resources. Good judgment will be required in selecting the

program issues most deserving intensive evaluation attention. In some instances modest inquiries will be sufficient to surface such deficiencies in structure and process that no expensive outcome evaluation is necessary. Further, evaluators will continue to be required to be ingenious in conducting powerful studies with modest financial resources.

Introduction of regulatory evaluation will hasten the professionalization of evaluation research. Standards for evaluation practice and credentials for evaluators will assume great importance. Certification and licensing might be required for regulatory evaluators. In light of the diversity within the field regarding methodological priorities, formulation of explicit standards for regulatory evaluation will provoke great controversies. Premature codification and enforcement of evaluation standards might commit the evaluation field to practice patterns which will constrain its long term development. (See, for example, Morell and Flaherty, 1978.) In various substantive sectors regulatory bodies will be challenged to choose wisely among competing methodological claims. They also will be well advised to be alert to possibilities for incorporating improvements in evaluation methodology.

For some researchers the explicit identification of evaluation with regulation will be troublesome. Some evaluation researchers prefer to see themselves as agents of program development rather than regulation. Many evaluators strive to avoid a regulatory identity because the obstacles it can create in securing needed cooperation from program personnel. In the long run conscientious regulatory evaluation should make service operators a good deal more serious about conducting effective

programs. The greater need to develop programs which will stand up under rigorous testing should stimulate demand for formative evaluation. Researchers who are not willing to work within a regulatory framework indirectly, then, may find their opportunities to do formative evaluation enhanced.

Evaluators who work within a regulatory framework will learn to contend with overt conflict with program personnel. They will come to expect to be greeted with some antagonism and learn to insulate themselves from it. Fortunately their link to program finances will put them in a position to demand information. To gain needed cooperation regulatory evaluators will not have to rely fully on the human relations techniques employed extensively by evaluators who are entirely dependent on voluntary cooperation.

Regulatory evaluation will expand substantially the range of evaluation findings made available to the public. Particularly for programs enjoying inflated reputations, publication of sound evaluation findings may lead the public to draw more drastic negative conclusions than are justified. A public unprepared for news of modest achievements may turn against once favored programs. The controversy triggered by the Westinghouse Head Start evaluation illustrates the problem (Evans, 1969). Regulatory bodies will be well advised to prepare the public for evaluation reports which may be particularly troublesome. Advance publicity about key evaluation questions and the methods may be useful. Responsible discussions of implications of findings should accompany publicity about negative results.

In the long run the publication issue may not be serious. Once accustomed to deflated accounts of programs the public may learn to respond without alarm to evaluation reports showing only modest program accomplishments. Ultimately it should be possible to develop a more sophisticated public which supports programs whose goals are modest but clearly attainable.

The proposed expansion of the methodological scope of regulatory agencies will only add to the importance of issues raised generally about regulatory agencies. A number of distortions are possible. Regulatory evaluation may come to be dominated by technicians for whom methodology becomes its own objective rather than a means of guiding public funding of social programs. Alternately program entrepreneurs may find ways of neutralizing the impact of regulatory evaluation either by gaining control over regulatory bodies or by undermining their political support. If regulatory evaluation is to make a useful contribution, it must reflect the legitimate concerns of the full social program constituency - including taxpayers, program operators, practitioners, clients, and evaluation researchers. Regulatory evaluation will do no more than provide evaluation researchers with greater leverage in conducting studies and influencing decisions. Considerable skill and energy will continue to be required to conduct evaluation research effectively even with the advantages of a regulatory framework. Evaluators will have to demonstrate their ability to contribute effectively as regulators if the model is to be institutionalized and extended.

REFERENCES

BERNSTEIN, I. and H.E. FREEMAN. (1975) Academic and Entrepreneurial Research. New York: Russell Sage.

CAMPBELL, D.T. (1969) "Reforms as Experiments." American Psychologist, 24: 409-429.

CAMPBELL, D.T. (1977) "Keeping the Data Honest in the Experimenting Society," in H.W. Melton and D. Watson (eds.) Interdisciplinary Dimensions of Accounting for Social Goals and Social Organizations. Columbus, Ohio: Grid, 37-76.

CARO, F.G. (1977) "Evaluation Research: An Overview." In F.G. Caro (ed.) Readings in Evaluation Research. Second edition. New York: Russell Sage.

CARO, F.G. (1977) "Experimental Methodology and Innovative Social Programming," in F.G. Caro (ed.) Readings in Evaluation Research. Second edition. New York: Russell Sage.

COOK, T.D. and C.L. GRUDER. (1978) "Metaevaluation Research." Evaluation Quarterly, 2, 5-51.

ETZIONI, A. (1960) "Two Approaches to Organization Analysis: A Critique and a Suggestion." Admin. Sc. Quarterly, 5, 257-278.

EVANS, J.W. (1969) "Head Start: Comments on the Criticism." Britannica Review of American Education, 1, 253-260.

FAIRWEATHER, G.W. (1964) Social Psychology in the Treatment of Mental Illness: An Experimental Approach. New York: Wiley.

FAIRWEATHER, G.W. et al. (1969) Community Life for the Mentally Ill. Chicago: Aldine.

GLASS, G.V. (1971) "The Growth of Evaluation Methodology." AERA Curriculum Evaluation Monograph Series, No. 7. Chicago: Rand McNally.

MORELL, J.A. and FLAHERTY, E.W. (1978) The development of evaluation as a profession: Current status and some predictions. Evaluation and Program Planning, 1, 11-17.

NEW YORK CITY OFFICE OF THE COMPTROLLER. (1978) Report on the Quality of Care and Operating Practices of the Home Attendant Program E78-429. New York: Office of the Comptroller, Bureau of Audit and Control.

RODMAN, H. and R. Kolodny. (1964) "Organizational Strains in the Researcher-Practitioner Relationship." In F.G. Caro (ed.) Readings in Evaluation Research. Second edition. New York: Russell Sage, 1977.

SCHULBERG, H.C. and F. BAKER. (1968) "Program Evaluation Models and the Implementation of Research Findings." In F.G. Caro (ed.) Readings in Evaluation Research. Second edition. New York: Russell Sage, 1977.

TORNATZKY, L.G. (1979) "The Triple Threat aluator." Evaluation and Program Planning, Vol. 2, No. 2, 111-115.

WIENK, R.E., C.E. REID, J.C. SIMONSON, F.J. EGGERS. (1979) Measuring Racial Discrimination in American Housing Markets: The Housing Market Practices Survey. Washington, D.C.: U.S. Department of Housing & Urban Development, Office of Policy Development & Research, Division of Evaluation.